ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
UIT TOÀN DIỆN • SÁNG TẠO • PHỤNG SỰ

BÀI GIẢNG
XÁC SUẤT THỐNG KÊ

Giảng viên: TS. PHÙNG MINH ĐỨC

(Bộ môn Toán Lý)

Nội dung môn học

1. Chương 1: Các khái niệm cơ bản về Xác suất
2. Chương 2: Biến ngẫu nhiên
3. Chương 3: Luật số lớn và định lý giới hạn trung tâm
4. Chương 4: Vectơ ngẫu nhiên
5. Chương 5: Thống kê toán học

Chương 1: Các khái niệm cơ bản về Xác suất

1.1 Phép đếm

Định lý 1.1.1 (Quy tắc nhân)

Cho một công việc được thực hiện bởi n giai đoạn. Giai đoạn thứ i có k_i cách thực hiện ($i = 1, 2, \dots, n$). Số cách thực hiện công việc trên là $k_1 k_2 \dots k_n$.

Ví dụ 1.1.2

An định mua một laptop và một headphone. Cửa hàng UITLaptop có 5 laptop và 6 headphone phù hợp với yêu cầu của An. Hỏi An có bao nhiêu cách chọn laptop và headphone?

Giải. Số cách chọn laptop và headphone là $5.6 = 30$ (cách).

Ví dụ 1.1.3

Có bao nhiêu số tự nhiên có 4 chữ số được lập từ các chữ số 1,2,3,4,5,6?

Giải.

- Cách chọn chữ số hàng nghìn: 6 cách
- Cách chọn chữ số hàng trăm: 6 cách
- Cách chọn chữ số hàng chục: 6 cách
- Cách chọn chữ số hàng đơn vị: 6 cách

Số các số thỏa yêu cầu là $6.6.6.6 = 1296$.

Định lý 1.1.4 (Quy tắc cộng)

Một công việc có thể chia thành k trường hợp để thực hiện

- Trường hợp 1 có n_1 cách thực hiện xong công việc
- Trường hợp 2 có n_2 cách thực hiện xong công việc
- $\ldots$
- Trường hợp k có n_k cách thực hiện xong công việc

và không có một cách thực hiện nào ở trường hợp này lại trùng với một cách thực hiện ở trường hợp khác. Khi đó, ta có $n_1 + n_2 + \ldots + n_k$ cách thực hiện xong công việc trên.

Ví dụ 1.1.5

Để đi từ A đến C có thể đi qua B, trong đó có 3 đường khác nhau đi trực tiếp từ A đến C, có 2 đường khác nhau để đi từ A đến B và có 4 đường khác nhau để đi từ B đến C. Hỏi có bao nhiêu cách đi từ A đến C?

Giải. Đi từ A đến C có 2 trường hợp:

- Đi trực tiếp từ A đến C : có 3 cách
- Đi từ A đến B sau đó đến C : có $2 \times 4 = 8$ (cách).

Số cách đi từ A đến C là $3 + 8 = 11$ (cách).

Ví dụ 1.1.6

Có bao nhiêu số tự nhiên chẵn có 3 chữ số khác nhau được tạo thành từ các chữ cố 0, 1, 2, 3, 4?

Giải. Trường hợp 1: Chữ số hàng trăm là chẵn.

- Chữ số hàng trăm: có 2 cách chọn
- Chữ số hàng đơn vị: có 2 cách chọn
- Chữ số hàng chục: có 3 cách chọn

Trường hợp 2: Chữ số hàng trăm là lẻ.

- Chữ số hàng trăm: có 2 cách chọn
- Chữ số hàng đơn vị: có 3 cách chọn
- Chữ số hàng chục: có 3 cách chọn

Số các số thỏa yêu cầu đề bài là $2.2.3 + 2.3.3 = 30.$

Định nghĩa 1.1.7 (Hoán vị)

Một hoán vị của n phần tử là một cách sắp xếp n phần tử đó theo một thứ tự nhất định.

Số hoán vị của n phần tử là $n!$.

Ví dụ 1.1.8

Cho 3 chữ cái A, B và C. Các hoán vị của A, B và C là $ABC, ACB, BAC, BCA, CAB, CBA$.

Định nghĩa 1.1.9 (Chỉnh hợp)

Một chỉnh hợp chập k của n phần tử là một nhóm **có thứ tự** gồm k phần tử **khác nhau** lấy từ n phần tử đã cho $(k \leq n)$.

Số chỉnh hợp chập k của n phần tử được kí hiệu và xác định như sau

$$\boxed{A_n^k = n.(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}}$$

Ví dụ 1.1.10

Một nhóm có 8 sinh viên cần chọn ra 1 người làm trưởng nhóm, 1 người làm thư ký và 1 người làm thủ quỹ. Hỏi có bao nhiêu cách chọn?

Định nghĩa 1.1.11 (Chỉnh hợp lặp)

Một chỉnh hợp lặp chập k của n phần tử là một nhóm **có thứ tự** gồm k phần tử **có thể giống nhau** lấy từ n phần tử đã cho.

Số chỉnh hợp lặp chập k của n phần tử được kí hiệu và xác định như sau

$$\boxed{\overline{A}_n^k = n^k.}$$

Ví dụ 1.1.12

Có thể lập được bao nhiêu số có 3 chữ số từ các chữ số 1,2,3,4,5?

Giải. Chọn 3 chữ số từ 5 chữ số có thứ tự và có thể giống nhau. Số các số thỏa yêu cầu đề bài là $5^3 = 125$.

Định nghĩa 1.1.13 (Tổ hợp)

Một tổ hợp chập k của n phần tử là một nhóm không phân biệt thứ tự gồm k phần tử khác nhau lấy từ n phần tử đã cho $(k \leq n)$.

Số tổ hợp chập k của n phần tử được kí hiệu và xác định như sau

$$\boxed{C_n^k = \frac{n!}{k!(n-k)!}.}$$

Ví dụ 1.1.14

Mỗi đề thi gồm 3 câu hỏi lấy trong 25 câu hỏi cho trước. Hỏi có thể lập được bao nhiêu đề thi có nội dung khác nhau?

1.2 Phép thử và biến cố

Định nghĩa 1.2.1

1. Việc thực hiện một nhóm các điều kiện nhất định để quan sát một hiện tượng nào đó được gọi là một **phép thử** (experiment).
2. Biến cố (sự kiện) **sơ cấp** là kết quả (outcome) của phép thử.
3. Tập hợp tất cả các biến cố sơ cấp của một phép thử được gọi là **không gian mẫu** (sample space), ký hiệu là Ω.
4. Một tập con của không gian mẫu được gọi là **sự kiện** hay **biến cố** (event). Ký hiệu các biến cố (sự kiện) là các chữ cái in hoa $A, B, C, ...$
5. Biến cố (sự kiện) **phức hợp** là biến cố (sự kiện) có thể phân tích thành các biến cố (sự kiện) nhỏ hơn.

Ví dụ 1.2.2

Tung một đồng xu là một phép thử. Kết quả sẽ xuất hiện mặt sấp hay ngửa (hình trái là mặt sấp, hình phải là mặt ngửa)

Mặt sấp - Mặt ngửa

- Kết quả "Xuất hiện mặt sấp" là một biến cố (sự kiện). Đặt S là biến cố "Xuất hiện mặt sấp", N là biến cố "Xuất hiện mặt ngửa".
- Không gian mẫu là $\{S, N\}$.

Ví dụ 1.2.3

Tung một con xúc xắc là một phép thử. Kết quả xuất hiện mặt 1, 2, 3, 4, 5 hay 6 chấm.

- Các kết quả: Xuất hiện mặt 1 chấm; xuất hiện mặt có số chấm lẻ; xuất hiện mặt có số chấm lớn hơn 4 là các biến cố (sự kiện).
- Xuất hiện mặt 1 chấm là biến cố sơ cấp. Xuất hiện mặt có số chấm lẻ là một biến cố phức hợp vì nó có thể phân tích thành các biến cố: Xuất hiện mặt 1 chấm, xuất hiện mặt 3 chấm, xuất hiện mặt 5 chấm.
- Đặt A_i là biến cố "Xuất hiện mặt i chấm" với $i = 1, 2, \ldots, 6$. Không gian mẫu là $\{A_1, A_2, A_3, A_4, A_5, A_6\}$.

Định nghĩa 1.2.4

1. Biến cố không thể xảy ra khi thực hiện phép thử được gọi là **biến cố không thể**, kí hiệu là $\varnothing$.
2. Biến cố luôn xảy ra khi thực hiện phép thử được gọi là **biến cố chắc chắn**, kí hiệu là Ω.
3. Biến cố **ngẫu nhiên** là biến cố có thể xảy ra hoặc không thể xảy ra.
4. Hai hay nhiều biến cố trong một phép thử có khả năng xảy ra như nhau được gọi là **đồng khả năng**.

Ví dụ 1.2.5

- Phép thử: Tung một con xúc xắc (cân đối và đồng chất) và xem số chấm.
- Biến cố không thể: Mặt có số chấm lớn hơn 6.
- Biến cố chắc chắn: Mặt có số chấm nhỏ hơn 7.
- Biến cố ngẫu nhiên: mặt có số chấm lẻ; mặt có số chấm nhỏ hơn 3.
- Các biến cố A_i : "Xuất hiện mặt i chấm" với $i = 1, 2, \ldots, 6$ là các biến cố đồng khả năng.

Định nghĩa 1.2.6

Cho A, B là các biến cố ngẫu nhiên.

1. Biến cố A được gọi là **thuận lợi** cho biến cố B, ký hiệu $A \subset B$, nếu A xảy ra thì B xảy ra.
2. **Biến cố tích** của A và B là biến cố khi A và B cùng xảy ra, kí hiệu $A \cap B$ hoặc AB.
3. Nếu $A \cap B = \varnothing$ thì A và B gọi là hai **biến cố xung khắc.**
4. **Biến cố tổng** của A và B là biến cố khi có ít nhất một trong hai biến cố A, B xảy ra, kí hiệu $A \cup B$. Nếu A, B là hai biến cố xung khắc thì biến cố tổng của A và B được kí hiệu là $A + B$.
5. **Hiệu** của hai biến cố A và B là biến cố khi A xảy ra nhưng B không xảy ra, kí hiệu $A \setminus B$.
6. Biến cố **đối** của biến cố $A : \overline{A} = \Omega \setminus A$ là biến cố xảy ra khi và chỉ khi A không xảy ra.

Ví dụ 1.2.7

Phép thử: gieo một con xúc xắc.

- Đặt A_i là biến cố "Xuất hiện mặt i chấm," với $i = 1, 2, \ldots, 6$.
- Nếu $i \neq j$ thì A_i và A_j là hai biến cố xung khắc.
- Đặt L là biến cố "Xuất hiện mặt có số chấm lẻ". Các biến cố A_1, A_3, A_5 là thuận lợi đối với L và $L = A_1 + A_3 + A_5$. Ngoài ra biến cố "Xuất hiện mặt có số chấm chẵn" = $\overline{L}$.

Một số tính chất:

Với các biến cố A, B:

- $\overline{A \cup B} = \overline{A} \cap \overline{B}; \quad \overline{A \cap B} = \overline{A} \cup \overline{B};$
- $A \cup B = A \cup (\overline{A} \cap B) = A \cup [B \setminus (A \cap B)]$.
- $A \cap \overline{A} = \varnothing; \quad A \cup \overline{A} = \Omega$.

Định nghĩa 1.2.8 (Hệ đầy đủ)

Một tập các biến cố $\{A_1, A_2, \ldots, A_n\}$ được gọi là một **hệ đầy đủ** nếu các biến cố này thỏa mãn hai điều kiện sau:

- Nếu $i \neq j$ thì A_i và A_j là hai biến cố xung khắc (tức là $A_i \cap A_j = \varnothing$)
- Phải có ít nhất một biến cố xảy ra, tức là

$$A_1 + A_2 + \cdots + A_n = \Omega.$$

Ví dụ 1.2.9

Cho A là một biến cố. Khi đó $\{A, \overline{A}\}$ là một hệ đầy đủ.

Nhận xét 1.2.10

1. Mọi biến cố (sự kiện) ngẫu nhiên đều có thể biểu diễn dưới dạng tổng của một số biến cố (sự kiện) sơ cấp nào đó.
2. Biến cố chắc chắn Ω là tổng của tất cả các biến cố (sự kiện) sơ cấp có thể.

Tính chất 1.2.12 Cho A, B và C là các biến cố của một phép thử. Khi đó

1. $A \cap B = B \cap A; A \cup B = B \cup A$
2. $A \cap (B \cap C) = (A \cap B) \cap C; (A \cup B) \cup C = A \cup (B \cup C)$
3. $A \cap \varnothing = \varnothing; A \cup \varnothing = A;$
   $A \cup \Omega = \Omega; A \cap \Omega = A$
4. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C); A \cup (B \cap C) = (A \cup B) \cap (A \cup C);$

1.3 Định nghĩa xác suất

Định nghĩa 1.3.1

Xác suất (probability) của một biến cố (sự kiện) A là một số nằm giữa 0 và 1, số này đo khả năng xuất hiện của biến cố (sự kiện) A khi thực hiện phép thử. Ký hiệu là $P(A)$.

Các định nghĩa xác suất

- Định nghĩa cổ điển
- Định nghĩa theo quan điểm hình học
- Định nghĩa theo quan điểm thống kê
- Định nghĩa theo tiên đề

Định nghĩa 1.3.2 (Định nghĩa xác suất cổ điển)

Xét một phép thử ngẫu nhiên có không gian mẫu Ω gồm hữu hạn phần tử và các biến cố sơ cấp là đồng khả năng. Giả sử A là một biến cố. Gọi

$$|A| = \text{số các kết quả thuận lợi cho } A,$$
$$|\Omega| = \text{số các phần tử của } \Omega.$$

Khi đó xác suất của A là

$$P(A) = \frac{|A|}{|\Omega|}.$$

Ví dụ 1.3.3 Xét phép thử là tung 1 xúc xắc và quan sát số chấm xuất hiện. Tính xác suất xuất hiện mặt có số chấm lớn hơn 3?

Giải:

- Gọi A_i là biến cố "Xuất hiện mặt i chấm" với $i = 1, 2, 3, \ldots, 6$.
- Không gian mẫu là $\Omega = \{A_1, A_2, A_3, A_4, A_5, A_6\} \Rightarrow |\Omega| = 6$.
- Gọi $B =$ "Xuất hiện mặt có số chấm lớn hơn 3" $= \{A_4, A_5, A_6\}$ $\Rightarrow |B| = 3$.

Do đó, xác suất của B là

$$P(B) = \frac{3}{6} = 0.5.$$

Ví dụ 1.3.4 Một hộp chứa 10 sản phẩm trong đó có 3 phế phẩm. Tính xác suất:

a. Chọn ngẫu nhiên 1 sản phẩm từ hộp được phế phẩm.

b. Chọn ngẫu nhiên 1 lần từ hộp ra 2 sản phẩm được 2 phế phẩm.

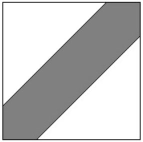
Ví dụ 1.3.5 Một hộp có 6 bi trắng, 4 bi đỏ và 2 bi đen. Chọn ngẫu nhiên 6 bi. Tìm xác suất để chọn được 3 bi trắng, 2 bi đỏ và 1 bi đen.

Nhận xét 1.3.6 Định nghĩa cổ điển về xác suất có ưu điểm là dễ vận dụng tuy nhiên định nghĩa này chỉ áp dụng được với các phép thử có hữu hạn kết quả đồng khả năng xảy ra.

Định nghĩa 1.3.7 (Định nghĩa theo quan điểm hình học)

Giả sử một điểm được rơi ngẫu nhiên vào một miền D và A là một mền con của D. Khi đó xác suất để điểm rơi ngẫu nhiên vào miền A được xác định bởi công thức:

$$P = \frac{\text{"độ đo" miền } A}{\text{"độ đo" miền } D}$$



Từ "độ đo" được hiểu như là **độ dài, diện tích, thể tích** tùy theo từng bài toán cụ thể.

Ví dụ 1.3.8 Cho một đoạn dây có độ dài 5cm. An dùng kéo cắt ngẫu nhiên tại một vị trí trên đoạn dây. Tính xác suất để An cắt được một đoạn dây có độ dài nhỏ hơn hoặc bằng 1cm.

Giải:

Gọi đoạn dây là MN và lấy 2 điểm P, Q trên MN sao cho $MP = 1cm$ và $QN = 1cm$.

Gọi A là biến cố "Cắt được một đoạn dây có độ dài nhỏ hơn hoặc bằng 1cm". Để biến cố A xảy ra, vị trí cắt phải thuộc MP hoặc QN. Do đó xác suất của biến cố A là

$$P(A) = \frac{2}{5} = 0,4.$$

Định nghĩa 1.3.10 (Định nghĩa theo quan điểm thống kê)

Thực hiện n phép thử độc lập với các điều kiện giống nhau có k lần xuất hiện biến cố A. Tỉ số

$$f(A) = \frac{k}{n}$$

được gọi là tần suất xuất hiện biến cố A trong n phép thử.

Định lý 1.3.11 (Luật số lớn Bernoulli)

Khi số phép thử n càng lớn thì tần suất xuất hiện biến cố A tiến về một giá trị xác định. Ta định nghĩa giá trị đó là xác suất của biến cố A.

Ví dụ 1.3.12 Để nghiên cứu khả năng xuất hiện mặt sấp khi tung một đồng xu, người ta tiến hành tung một đồng xu nhiều lần (đồng xu không cần cân đối đồng chất nhưng các lần tung phải giống nhau) và thu được kết quả sau đây:

Người làm TN	n	k	Tần suất $\frac{k}{n}$
Buffon	4040	2048	0,5069
Pearson	12000	6019	0,5016
Pearson	24000	12012	0,5005

Ta thấy khi số phép thử tăng lên thì tần suất xuất hiện mặt sấp dao động quanh giá trị 0,5. Điều này cho phép ta hy vọng rằng khi số phép thử tăng lên vô hạn thì tần suất xuất hiện mặt sấp hội tụ về 0,5.

Nhận xét 1.3.13 Từ định nghĩa này trong thống kê, người ta hay dùng khái niệm tỷ lệ thay cho xác suất. Chẳng hạn tỷ lệ hạt thóc nảy mầm trong cùng một điều kiện về môi trường là 60%, nghĩa là khi chọn một hạt thóc ngẫu nhiên thì xác suất của biến cố A hạt thóc nảy mầm là 0,6 hay $P(A) = 0,6$.

Nhận xét 1.3.14 Định nghĩa xác suất theo quan điểm thống kê của xác suất khắc phục được một nhược điểm của định nghĩa cổ điển là không dùng đến khái niệm đồng khả năng. Nó hoàn toàn dựa trên các quan sát thực tế để làm cơ sở kết luận về xác suất xảy ra của một biến cố

Chỉ áp dụng được đối với các hiện tượng ngẫu nhiên mà tần suất của nó có tính ổn định. Hơn nữa, để xác định một cách tương đối chính xác giá trị của xác suất, ta phải tiến hành trên thực tế một số lượng đủ lớn các phép thử.

Định nghĩa 1.3.15 (Định nghĩa xác suất theo hệ tiên đề)

Xét phép thử có không gian mẫu là Ω. Một ánh xạ $P : \Omega \to [0, 1]$ được gọi là hàm xác suất trên Ω nếu nó thỏa mãn các điều kiện sau:

1. Với mỗi biến cố A ta có $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. Với bất kì họ đếm được các biến cố đôi một xung khắc $A_1, A_2, \ldots$ ta có:

$$P(A_1 \cup A_2 \cup \ldots) = \sum_{i \geq 1} P(A_i)$$

Năm 1933, nhà toán học Andrey Nikolaevich Kolmogorov (1903 – 1987) đã đưa ra định nghĩa xác suất theo phương pháp tiên đề. Ông là một nhà toán học Liên Xô đã có nhiều đóng góp lớn trong lý thuyết xác suất và tôpô.

Nguồn: https://www.kolmogorov.com/Kolmogorov.html

Tính chất 1.3.16 Xét không gian mẫu Ω và các biến cố A, B, C bất kì. Ta có

1. $0 \leq P(A) \leq 1$
2. $P(\varnothing) = 0; P(\Omega) = 1$
3. Nếu $A \cap B = \varnothing$ thì $P(A \cup B) = P(A) + P(B)$
4. $B \subset A \Rightarrow P(A \setminus B) = P(A) - P(B)$
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
6. $P(\overline{A}) = 1 - P(A)$
7. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$

Ví dụ 1.3.17 Trong một lớp chuyên tin học, người ta thấy rằng có 28% học sinh thích ngành CNPM, 15% học sinh thích ngành KHMT và 8% học sinh thích cả hai ngành trên. Tính xác suất để gặp một học sinh không thích ngành KHMT và cũng không thích ngành CNPM.

Giải. Đặt $A =$ gặp học sinh thích ngành CNPM, $B =$ gặp học sinh thích ngành KHMT. Ta có $P(A) = 0,28$ và $P(B) = 0,15$ và $P(A \cap B) = 0,08$.

Xác suất để gặp một học sinh không thích ngành KHMT và ngành CNPM là

$$P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B).$$

Ta có:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,28 + 0,15 - 0,08 = 0,35.$$

Do đó xác suất để gặp một học sinh không thích ngành KHMT và cũng không thích ngành CNPM là

$$1 - 0,35 = 0,65.$$

Ví dụ 1.3.18 Một công ty phần mềm đang tuyển dụng các ứng viên cho 4 vị trí quan trọng trong ban quản lý văn phòng mới của họ ở Thủ Đức. Có 5 ứng cử viên đến từ Việt Nam và 3 đến từ Nhật. Xác suất để ít nhất 1 người Nhật được chọn là bao nhiêu?

Giải. Đặt A là biến cố "Có ít nhất một người Nhật" thì $\overline{A}$ là biến cố "Không có người Nhật nào", tức là cả 4 người đều là người Việt Nam.

Ta có

$$P(\overline{A}) = \frac{C_5^4}{C_8^4} = \frac{1}{14}$$

và do đó xác suất cần tìm là

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{1}{14} = \frac{13}{14}$$

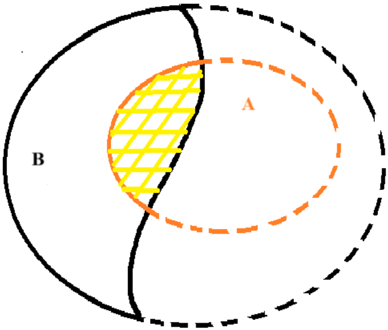
Ví dụ 1.3.19 Một lô hàng gồm 10 sản phẩm, trong đó có 2 phế phẩm. Lấy ngẫu nhiên đồng thời 6 sản phẩm từ lô hàng. Tính xác suất để có không quá 1 phế phẩn trong 6 sản phẩm lấy ra. (ĐS: 2/3)

Ví dụ 1.3.20 Một lớp học có 20 học sinh trong đó có 10 học sinh giỏi toán, 8 học sinh giỏi văn và 6 học sinh giỏi cả toán và văn. Chọn ngẫu nhiên một học sinh.

a. Tính xác suất để học sinh này giỏi ít nhất một môn. (ĐS: 0,6)

b. Tính xác suất để học sinh này không giỏi môn nào cả. (ĐS: 0,4)

1.4 Xác suất có điều kiện. Công thức nhân xác suất



Định nghĩa 1.4.1 (xác suất có điều kiện)

Giả sử trong một phép thử, ta có $P(B) > 0$. Khi đó xác suất có điều kiện của biến cố A với điều kiện đã có biến cố B (gọi là xác suất của A với điều kiện B), ký hiệu là $P(A|B)$, xác định như sau

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Ví dụ 1.4.2 Một lớp học có 56 sinh viên (SV), trong đó có 26 nam và 30 nữ. Kết quả thi môn XSTK có 7 SV nam và 5 SV nữ đạt điểm giỏi. Chọn ngẫu nhiên một sinh viên trong lớp.
a. Tính xác suất để chọn được sinh viên đạt điểm giỏi.
b. Tính lại xác suất để chọn được sinh viên đạt điểm giỏi biết rằng sinh viên đó là nữ.

Giải. a. Gọi A là biến cố "Chọn sinh viên đạt điểm giỏi". Khi đó

$$P(A) = \frac{12}{56}.$$

b. Đặt B là biến cố "Chọn sinh viên nữ", khi đó Biến cố $A \cap B$ là "Chọn sinh viên đạt điểm giỏi và là nữ". Ta có

$$P(B) = \frac{30}{56};\ P(A \cap B) = \frac{5}{56}.$$

Như vậy xác suất để chọn được sinh viên đạt điểm giỏi biết rằng sinh viên đó là nữ là

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{5}{30} = 0,167.$$

Định nghĩa 1.4.3 (Biến cố độc lập)

Hai biến cố A và B được gọi là **độc lập** nếu

$$P(A \cap B) = P(A).P(B).$$

Hay ta nói hai biến cố A, B được gọi là độc lập nếu xác suất của biến cố A không ảnh hưởng đến xác suất của biến cố B và ngược lại.

Định lý 1.4.4

Cho A, B là các biến cố của một phép thử. Các điều sau đây là tương đương:

1. A và B độc lập (tức là $P(A \cap B) = P(A)P(B)$);
2. $P(B|A) = P(B)$; $P(A|B) = P(A)$.

Mệnh đề 1.4.5

Cho A, B là hai biến cố độc lập. Khi đó

1. $\overline{A}$ và B là hai biến cố độc lập.
2. A và $\overline{B}$ là hai biến cố độc lập.
3. $\overline{A}$ và $\overline{B}$ là hai biến cố độc lập.

Ví dụ 1.4.6 Một thành phố có 51% nam và 49% nữ, và tỷ lệ nam và nữ mù màu được thể hiện trong bảng xác suất dưới đây:

	Nam	Nữ	Tổng cộng
Mù màu	0,04	0,002	0,042
Không mù màu	0,47	0,488	0,958
Tổng	0,51	0,49	1,00

a. Gặp ngẫu nhiên một người nam. Tính xác suất người này bị mù màu.
b. Biến cố gặp người nam và biến cố gặp người bị mù màu có độc lập không?
Giải. a. Đặt A là biến cố "Gặp người nam" và B là biến cố "Gặp người bị mù màu". Ta cần tính $P(B|A)$.
Ta đã có $P(A) = 0,51$ và $P(A \cap B) = 0,04$. Do đó

$$P(B|A) = \frac{0,04}{0,51} = 0,078.$$

b. Ta có $P(B) = 0,042 \neq P(B|A)$ nên A, B không độc lập.

Mệnh đề 1.4.7

Cho $A_1, A_2, \ldots, A_n$ là các biến cố bất kì.

1. Công thức nhân xác suất $P(A_1 \cap A_2) = P(A_1)P(A_2|A_1)$
2. $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap \ldots \cap A_{n-1})$
3. Nếu $A_1, A_2, \ldots, A_n$ là các biến cố đôi một độc lập nhau thì

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2) \ldots P(A_n).$$

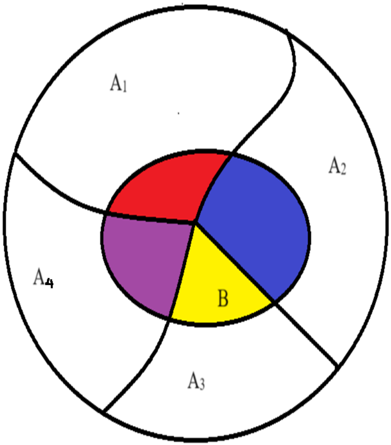
Ví dụ 1.4.8 Một tổ có 15 sinh viên trong đó có 5 sinh viên học giỏi môn Xác suất thống kê. Chia tổ này thành 5 nhóm, mỗi nhóm 3 người. Tính xác suất để nhóm nào cũng có một sinh viên học giỏi môn Xác suất thống kê.

Giải. Gọi A là biến cố "nhóm nào cũng có một sinh viên học giỏi môn Xác suất thống kê";
A_i là biến cố "Nhóm i có một sinh viên học giỏi môn Xác suất thống kê", $i = 1, \ldots, 5$.
Khi đó $A = A_1A_2A_3A_4A_5$. Sử dụng công thức nhân

$$P(A) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)P(A_4|A_1A_2A_3)P(A_5|A_1A_2A_3A_4)$$

$$= \frac{81}{1001}.$$

1.5 Xác suất toàn phần. Công thức Bayes



Cho hệ đầy đủ các biến cố $\{A_1, A_2, \ldots, A_n\}$ của một phép thử, tức là hệ thỏa mãn hai điều kiện sau:

- Nếu $i \neq j$ thì $A_i \cap A_j = \varnothing$;
- Phải có ít nhất một biến cố xảy ra, tức là

$$A_1 + A_2 + \cdots + A_n = \Omega,$$

và B là một biến cố bất kì. Khi đó, ta có

$$B = A_1B + A_2B + \cdots + A_nB.$$

Định lý 1.5.1 (Công thức xác suất đầy đủ/toàn phần)

$$\begin{aligned} P(B) &= P(A_1B) + P(A_2B) + \cdots + P(A_nB) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) \ldots + P(A_n)P(B|A_n). \end{aligned}$$

Từ công thức xác suất có điều kiện, ta có

$$P(A_kB) = P(A_k)P(B|A_k) = P(B)P(A_k|B), k = 1, 2, \ldots, n,$$

do đó ta có kết quả sau:

Định lý 1.5.2 (Công thức Bayes)

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{P(B)}, k = 1, 2, \ldots, n.$$

Thomas Bayes (Nguồn: Internet)

Thomas Bayes (1701-1761) là nhà thống kê học, nhà triết học người Anh.

Ví dụ 1.5.3 Trong một thành phố nọ, người ta thấy rằng tỷ lệ người hút thuốc lá là 30% và số người bị viêm phổi trong số người hút thuốc lá là 60%, tỷ lệ người viêm phổi trong số người không hút thuốc là 40%.

a. Chọn ngẫu nhiên 1 người. Tính xác suất để người đó bị viêm phổi.

b. Chọn ngẫu nhiên 1 người, biết rằng người đó viêm phổi. Tính xác suất người đó hút thuốc lá.

Giải.

Gọi A là biến cố 'Chọn ra một người bị viêm phổi".

Gọi B là biến cố "Người được chọn ra là người hút thuốc", ta có $\overline{B}$ là biến cố "Người được chọn ra là người không hút thuốc".

Hệ biến cố đầy đủ là $\{B, \overline{B}\}$.

Ta có $P(B) = 0,3; P(\overline{B}) = 0,7$ và $P(A|B) = 0,6; P(A|\overline{B}) = 0,4$.

Do đó: ....

....

Ví dụ 1.5.4 Hai nhà máy cùng sản xuất một loại sản phẩm. Tỉ lệ sản phẩm tốt của nhà máy thứ nhất là 90%; của nhà máy thứ hai là 85%. Từ một kho chứa 1/3 sản phẩm của nhà máy thứ nhất và 2/3 sản phẩm còn lại của nhà máy thứ hai, người ta lấy ra 1 sản phẩm để kiểm tra.
a. Tính xác suất để lấy được sản phẩm không tốt.
b. Giả sử sản phẩm lấy là sản phẩm tốt. Tính xác suất sản phẩm đó do nhà máy thứ 2 sản xuất.
Giải. Đặt A_i là "Sản phẩm do nhà máy i sản suất" với $i = 1, 2$.
Khi đó A_1, A_2 là một hệ đầy đủ.
Đặt A là "Sản phẩm không tốt".

....

....

....

....

Bài tập 1.5.5. Có hai lô hàng: lô thứ nhất có 7 sản phẩm loại A và 3 sản phẩm loại B; lô thứ hai có 8 sản phẩm loại A và 4 sản phẩm loại B.

a. Lấy ngẫu nhiên một lô hàng, rồi từ lô hàng đó lấy ra 2 sản phẩm. Tính xác suất trong 2 sản phẩm lấy ra có loại A.

b. Nếu 2 sản phẩm lấy ra có loại A thì xác suất để các sản phẩm loại A đó thuộc lô thứ 2 là bao nhiêu?